AD-A049 556 MASSACHUSETTS INST OF TECH CAMBRIDGE DEPT OF MATHEMATICS F/G 12/1

APPROXIMATE SOLUTIONS FOR CERTAIN OPTIMAL STOPPING PROBLEMS.(U)

DEC 77 A J PETKAU N00014-75-C-0555

UNCLASSIFIED TR-8 NL

1 OF 1

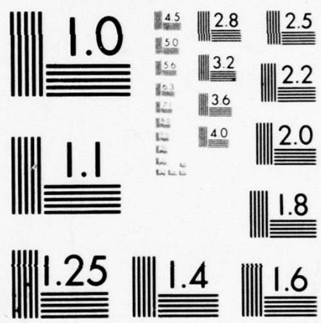
ADA049556

END

DATE FILMED

3-78

DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A049556

AD No.
DDC FILE COPY

12

APPROXIMATE SOLUTIONS FOR CERTAIN OPTIMAL STOPPING PROBLEMS

BY

A. JOHN PETKAU
(University of British Columbia, CANADA)

TECHNICAL REPORT NO. 8

DECEMBER 21, 1977

DDC
FEB 3 1978
F

PREPARED UNDER CONTRACT
N00014-75-C-0555 (NR-042-331)
FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF MATHEMATICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

6 APPROXIMATE SOLUTIONS FOR CERTAIN OPTIMAL STOPPING PROBLEMS.

by

10 A. John Petkau
(University of British Columbia, CANADA)

9 Technical Report No. 8

11 21 Dec[redacted]77

12 26 p.

14 TR-8

Prepared Under Contract
15 N00014-75-C-0555, (NR-042-331)
For the Office of Naval Research

Reproduction in Whole or in Part is Permitted
for any purpose of the United States Government

Approved for public release; distribution unlimited

DEPARTMENT OF MATHEMATICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS

This report will also appear as a Stanford Technical Report under
U.S. Army Grant DAAG29-77-G-0031.

220 021

ACCESSION for
NTIS
DDC
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. SPECIAL
A

Approximate Solutions for Certain Optimal Stopping Problems

A. John Petkau

1. INTRODUCTION

The following optimal stopping problem (which is one of several different problems which have come to be known as the one-armed bandit problem) has arisen both in the study of a sequential sampling inspection plan by Chernoff and Ray (1965) and in the study of a sequential model for clinical trials by Chernoff (1967) as well as in connection with problems involving the sequential estimation of the common mean of two normal populations considered by Mallik (1971) and Chernoff (1971): Let $X(t)$ be a Wiener process described by $E[dX(t)] = \mu dt$ and $\mathrm{Var}[dX(t)] = \sigma^2 dt$ where σ^2 is presumed known. One is permitted to stop observing the process $X(t)$ at any time t, $0 \le t \le N$, and receive a payoff $X(t)$. The unknown parameter μ is assumed to have a $N(\mu_0, \sigma_0^2)$ prior. What is the optimal stopping procedure?

It is easy to verify that the posterior distribution of μ given $X(t')$, $0 \le t' \le t$, is $N(Y^*(s^*), s^*)$ where

$$Y^*(s^*) = \frac{\mu_0\sigma_0^{-2} + X(t)\sigma^{-2}}{\sigma^{-2} + t\sigma^{-2}} \tag{1.1}$$

and

$$s^* = (\sigma_0^{-2} + t\sigma^{-2})^{-1} . \tag{1.2}$$

Here s^* varies from $s_0^* = \sigma_0^2$ to $s_1^* = (\sigma_0^{-2} + N\sigma^{-2})^{-1}$. Furthermore, the process $Y^*(s^*)$ is a Wiener process (in the $-s^*$ scale) described by $E[dY^*(s^*)] = 0$ and $Var[dY^*(s^*)] = -ds^*$, starting from $Y^*(s_0^*) = \mu_0$. The loss upon stopping at $(Y^*(s^*), s^*)$ is $-X(t)$ which from (1.1) is a linear function of $-Y^*(s^*)/s^*$. Applying the transformation $s = s^*/s_1^*$, $Y(s) = Y^*(s^*)/s^{*1/2}$ leads to a normalized version of this stopping problem where s varies from $s_0 = \sigma_0^2(\sigma_0^{-2} + N\sigma^{-2})$ to $s_1 = 1$ and in which the stopping cost is given by $d(y,s)$ defined by

$$d(y,s) = -y/s \tag{1.3}$$

for $s \geq 1$ with stopping enforced at $s = 1$.

This normalized problem is a special case of the following optimal stopping problem: Given a Wiener process $\{Y(s),\ s \geq s_1\}$ in the $-s$ scale described by $E[dY(s)] = 0$ and $Var[dY(s)] = -ds$ and starting at $Y(s_0) = y_0$, find the stopping time S to minimize $E[d(Y(S),S)]$. If we define $\rho(y_0,s_0) = \inf b(y_0,s_0)$ where $b(y_0,s_0)$ is the risk associated with a particular stopping time and the infimum is taken over all such stopping times, $\rho(y,s)$ represents the best that can be achieved once (y,s) has been reached, irrespective of how it was reached. An optimal procedure is then described by the continuation set $C = \{(y,s): \rho(y,s) < d(y,s)\}$. Chernoff (1968) has demonstrated that one should expect the solution (ρ, C) of the stopping problem to be a solution of the following free boundary problem (∂C denotes the boundary of the set C):

$$\begin{aligned} \tfrac{1}{2}\rho_{yy}(y,s) &= \rho_s(y,s) \quad \text{for } (y,s) \in C, \\ \rho(y,s) &= d(y,s) \quad \text{for } (y,s) \in C^c, \\ \rho_y(y,s) &= d_y(y,s) \quad \text{for } (y,s) \in \partial C. \end{aligned} \tag{1.4}$$

Furthermore, for any such stopping problem, Van Moerbeke (1974) has shown that one should never stop at points (y,s) at which $\frac{1}{2} d_{yy}(y,s) - d_s(y,s) < 0$.

Applying this criterion to the normalized version of the stopping problem described above, hereinafter referred to as the <u>one-armed bandit problem</u>, one finds that $\{(y,s): y > 0, s > 1\}$ must be a subset of the optimal continuation region $\mathcal{C}$. Chernoff and Ray (1965) have shown that for this problem $\mathcal{C}$ can be described as $\mathcal{C} = \{(y,s): y > \tilde{y}(s), s > 1\}$ and have determined asymptotic expansions for the boundary curve $\tilde{y}(s)$ in the regions of large s and s close to 1. The leading terms in these expansions are given by

$$\tilde{y}(s) \approx -(2s \log s)^{1/2} \quad \text{as} \quad s \to \infty$$

$$\tilde{y}(s) \approx -0.64(s-1)^{1/2} \quad \text{as} \quad s \to 1 .$$

The scale $z = y/s$ and $t = 1/s$ is more appropriate for applications and these expansions are illustrated in this scale in Chernoff and Ray (1965). These expansions also appear as the curves $\tilde{z}_0$ and $\tilde{z}_1$ in Figure C of this paper. It is evident from this illustration that these asymptotic expansions are inadequate as a complete description of the optimal continuation region. An approximation to the optimal continuation region is required as a description of the optimal procedure in the region where the asymptotic expansions are clearly inadequate.

Although it is possible that refined methods of asymptotic analysis could lead to expansions which would provide an adequate description of the optimal procedure, the purpose of the present paper is to describe simple methods which lead to arbitrarily accurate numerical approximations to the optimal continuation region for the one-armed bandit problem. Although most of the discussion in the present paper will concentrate on the one-armed bandit problem, these same methods could be applied with equal facility to any optimal stopping problem of the general form described above.

2. APPROXIMATE SOLUTIONS

In this section we indicate how approximate solutions to the one-armed bandit problem can be obtained by replacing the problem for the Wiener process $Y(s)$ by an analogous problem for a discrete-time discrete-step process which we will denote by $Y'(s')$.

Consider the process $Y'(s')$ which starts at $Y'(1+n\cdot\Delta) = y'$ and is defined by $Y'(s'-\Delta) = Y'(s') \pm \Delta^{1/2}$ each with probability $\frac{1}{2}$. This process is observed for at most n successive times and the cost associated with stopping the process at any point (y',s') is given by $d(y',s')$ defined by (1.3). The problem is to find a stopping time to minimize the expected loss. We shall denote the optimal expected loss by $\rho'(y',s')$. For this problem, a backward induction algorithm becomes

$$\begin{aligned}\rho'(y',1+n\cdot\Delta) = \min\{d(y',1+n\cdot\Delta),\ &\tfrac{1}{2}[\rho'(y'+\Delta^{1/2},1+(n-1)\cdot\Delta)\\ &+ \rho'(y'-\Delta^{1/2},1+(n-1)\cdot\Delta)]\}\end{aligned} \tag{2.1}$$

for $n > 1$ with $\rho'(y',1) = d(y',1)$. Using the methods of Chernoff and Petkau (1976), it is easy to verify that for this problem the optimal stopping set can be described as $\{(y',1+n\cdot\Delta): y' \le \tilde{y}_n(\Delta),\ n \ge 1\}$ where for each fixed value of Δ , $\{\tilde{y}_n(\Delta);\ n=1,2,...\}$ is a non-increasing non-positive sequence. Note that this set does not depend upon the initial point. Further note that direct application of (2.1) yields $\tilde{y}_1(\Delta) \equiv 0$.

Since $Y'(s')$ is a process of independent increments with mean zero and variance one per unit change in $-s'$, <u>any</u> stopping problem for the Wiener process $Y(s)$ of the previous section can be imitated by the use of a small value of Δ in the $Y'(s')$ process. As Δ approaches zero, the solution of the analogous discrete problem would be <u>expected</u> to converge to the solution of the Wiener process problem. In particular, for the one-armed bandit problem this leads to the initial approximation

$$\tilde{y}(1+n\cdot\Delta) \approx \tilde{y}_n(\Delta) \tag{2.2}$$

where $\tilde{y}(1+n\cdot\Delta)$ denotes the optimal boundary for the one-armed bandit problem evaluated at $s = 1+n\cdot\Delta$.

It remains to evaluate the sequence $\{\tilde{y}_n(\Delta)\}$. Consider the $Y'(s')$ process defined as above on the grid of points $\{(y',s'): s' = 1+n\cdot\Delta, y' = c + k\cdot\Delta^{1/2}; n = 0,1,2,\ldots,k=0, \pm 1, \pm 2,\ldots\}$. Note that the grid is completely specified by the parameter c (for convenience, assume $0 \leq c < \Delta^{1/2}$). Examination of (2.1) with the particular form of $d(y,s)$ given in (1.3) reveals that for any given choice of Δ, if the points $\{(y',1+n\cdot\Delta): y' \leq y^*\}$ are stopping points then so are the points $\{(y',1+(n+1)\cdot\Delta): y' \leq y^*-\Delta^{1/2}\}$. This observation, together with the fact that the sequence $\{\tilde{y}_n(\Delta)\}$ is non-increasing, implies that when using the backward induction algorithm (2.1) to classify the grid points as either stopping or continuation points, the comparisons implied by (2.1) need be carried out at only a single value of y' for each fixed value of s' . The algorithm (2.1) can now be easily implemented in a direct fashion.

Due to the special nature of the one-armed bandit problem, namely the fact that all points (y,s) with $y > 0$ and $s > 1$ are continuation points, one might expect to be able to improve somewhat upon the naive approach outlined above. Consider a particular path of the $Y'(s')$ process originating at the point $(y',s') = (c+2\cdot\Delta^{1/2}, 1+n\cdot\Delta)$. The path of the $Y'(s')$ process could hit the line $y' = c+\Delta^{1/2}$ for the first time at $s' = 1+(n-1)\cdot\Delta, 1+(n-3)\cdot\Delta, \cdots$. Alternately, the path could remain above the line $y' = c+\Delta^{1/2}$ all the way to $s' = 1$. Noting that the points $(c+\Delta^{1/2}, s')$ are continuation points for all $s' > 1$ (since $\tilde{y}_1(\Delta) = 0$ and the sequence $\tilde{y}_n(\Delta)$ is non-increasing) leads to the relation

$$\rho'(c+2\cdot\Delta^{1/2}, 1+n\cdot\Delta) = \sum_{m=1}^{n} p_m \rho'(c+\Delta^{1/2}, 1+(n-m)\cdot\Delta)$$

$$+ \sum_{k=1}^{n+1} q_{n,k} d(c+(k+1)\cdot\Delta^{1/2}, 1)$$

where p_m is the probability that an ordinary random walk starting at 0 first passes through 1 at time m and $q_{n,k}$ is the probability that an ordinary random walk starting at 0 stays above -1 until time n and achieves level $k-1$ at time n. From Feller (1968, p. 89, Theorem 2) one finds

$$\begin{aligned} p_m &= 0 && \text{for } m \text{ even},\\ &= \frac{1}{m}\cdot\binom{m}{\frac{m+1}{2}}\cdot 2^{-m} && \text{for } m \text{ odd}. \end{aligned} \tag{2.4}$$

In addition we have the recursive relation $p_{m+2} = \frac{m}{m+3} p_m$ with $p_1 = \frac{1}{2}$ and $p_2 = 0$. From Feller (1968, p. 73, Ballot Theorem) one also finds that

$$\begin{aligned} q_{n,k} &= 0 && \text{for } n \text{ even and } k \text{ odd ,} \\ &= 0 && \text{for } n \text{ odd and } k \text{ even ,} \\ &= \frac{k}{n+1} \cdot \binom{n+1}{\frac{n+k+1}{2}} \cdot 2^{-n} && \text{otherwise .} \end{aligned} \tag{2.5}$$

The relation (2.3) provides a modified method of carrying out the backward induction which we shall call the <u>truncation method</u>: At $s' = 1$, the risks are specified by $d(y,s)$. At any stage $s' = 1+n\cdot\Delta$, compute the risk at $y' = c+2\cdot\Delta^{1/2}$ by means of (2.3). The risks at the levels $y' = c+k\cdot\Delta^{1/2}$ for $k = 1,0,-1,-2,\ldots$ are computed using the algorithm (2.1) in the fashion described above.

Returning for a moment to the one-armed bandit problem, it is well-known (and obvious from (1.4)) that changing the stopping cost function $d(y,s)$ by adding to it any solution of the heat equation leaves the optimal continuation region unchanged. For the present purposes, it is convenient to consider the new stopping cost function $d_0(y,s)$ defined by $d_0(y,s) = d(y,s) + y$, that is,

$$d_0(y,s) = y - y/s \tag{2.6}$$

for $s \geq 1$ with stopping enforced at $s = 1$. Note that $d_0(y,1) \equiv 0$. Denoting the corresponding optimal risk by $\rho_0'(y,s)$, the algorithm (2.1) becomes

$$\rho_0'(y',1+n\cdot\Delta) = \min\{d_0(y',1+n\cdot\Delta),\ \tfrac{1}{2}[\rho_0'(y'+\Delta^{1/2},1+(n-1)\cdot\Delta) + \rho_0'(y'-\Delta^{1/2},1+(n-1)\cdot\Delta)]\} \tag{2.7}$$

for $n > 1$ with $\rho_0'(y',1) \equiv 0$. Further, the relation (2.3) becomes

$$\rho_0'(c+2\cdot\Delta^{1/2},1+n\cdot\Delta) = \sum_{m=1}^{n} p_m\cdot\rho_0'(c+\Delta^{1/2},1+(n-m)\cdot\Delta)\ . \tag{2.8}$$

This reduces the computation involved in carrying out the truncation method.

To this point we have simply described the direct and truncation methods of carrying out the backward induction algorithm for the $Y'(s')$ process when the motion of the process is restricted to a grid specified by a particular value of the parameter c.

It should be emphasized that carrying out the backward induction algorithm for a particular grid simply classifies all the grid points as either stopping or continuation points. The sequence $\{\tilde{y}_n(\Delta)\}$ itself is not determined. At each fixed value of $s' = 1+n\cdot\Delta$, the algorithm simply determines the two adjacent grid points between which the number $\tilde{y}_n(\Delta)$ must lie; that is, the algorithm determines the value of k $(= k(n,\Delta,c))$ such that

$$c + k\cdot\Delta^{1/2} \leq \tilde{y}_n(\Delta) < c + (k+1)\cdot\Delta^{1/2}\ .$$

However, implementing the algorithm for different grids all with the same fixed value of Δ but specified by a sequence of values of the parameter c between 0 and $\Delta^{1/2}$ allows the sequence $\{\tilde{y}_n(\Delta)\}$ to be determined to within any desired degree of accuracy. This will be indicated in more detail in the next section where the appropriate computations for the one-armed bandit problem are described.

To this point we have presented simple methods of obtaining heuristic initial approximations to the solutions of optimal stopping problems for a zero drift standard Wiener process. These methods involve replacing the Wiener process problem by an analogous discrete problem involving dichotomous random variables. The relation of the solution of any such Wiener process problem to the solutions of an entire class of analogous discrete problems is considered in Chernoff and Petkau (1976). A particular result of that paper is the following simple approximate relation between the optimal boundary of any such Wiener process problem and the optimal boundary of the corresponding analogous discrete problem for the $Y'(s')$ process described in the previous section:

$$\tilde{y}(1+n\cdot\Delta) = \tilde{y}_n(\Delta) \pm 0.5\,\Delta^{1/2} \tag{2.9}$$

(the sign being determined so as to make the continuation region for the Wiener process problem larger). For the one-armed bandit problem, this leads to the following refinement of (2.2)

$$\tilde{y}(1+n\cdot\Delta) \simeq \tilde{y}_n(\Delta) - 0.5\,\Delta^{1/2}\ . \tag{2.10}$$

This refinement takes the form of a "correction for continuity", the solution of the analogous discrete problem being corrected in order to obtain (approximately) the solution of the Wiener process problem. The behavior of the approximations (2.2) and (2.10) will be examined in the next section.

3. APPLICATIONS

To illustrate the accuracy of these approximations, it would be desirable to evaluate these approximations in a Wiener process problem for which the exact solution is known. A normalized version of a gambling problem discussed in Van Moerbeke (1974) is the following: Suppose $\{X(s);\ 0 < s \le 1\}$ is a zero drift standard Wiener process and that a gambler wins money at a constant rate whenever the process occupies the positive x half-plane and loses money at the same constant rate whenever the process occupies the negative x half-plane. If the gambler is permitted to stop the process at any time s, $0 < s \le 1$, what is the gambler's optimal strategy?

This problem can be formulated as an optimal stopping problem by defining the reward $g(x,s)$ for stopping the process at the point (x,s) to be

$$\begin{aligned} g(x,s) &= 1 - s + 2x^2 \qquad \text{for } x \ge 0 , \\ &= 1 - s \qquad \text{for } x \le 0 , \end{aligned}$$

the problem being to find a stopping time that maximizes the expected reward. Van Moerbeke (1974) proves that the optimal continuation region for this problem can be described as $\{(x,s): x > \tilde{x}(s),\ 0 < s \le 1\}$

where $\tilde{x}(s) = -\alpha(1-s)^{1/2}$ and α is the solution of the simple equation $\alpha \cdot \int_0^\infty \exp[\lambda\alpha-\lambda^2/2]d\lambda = 1$ which can easily be determined to be $\alpha \doteq 0.5061$. Modifying the reward function to be $\hat{g}(x,s) = g(x,s)-2[x^2+1-s]$ does not change the solution of this problem and it is easily seen that the methods of the previous section are directly applicable (in particular, $\hat{g}(x,1) \equiv 0$ for $x \geq 0$) .

This Wiener process problem has been approximated by three different analogous discrete problems, those corresponding to $\Delta = 0.01$, 0.0025 and 0.000625. For each fixed value of Δ the computations were carried out for all grids specified by values of the parameter c varying from 0 to $\Delta^{1/2}$ in steps of 0.001 . Thus each individual member of each of three sequences $\{\tilde{x}_n(\Delta)\}$ *is located* to within an error of 0.001. In addition the corrected sequences $\{\tilde{x}^*_n(\Delta)\}$ defined by $\tilde{x}^*_n(\Delta) = \tilde{x}_n(\Delta)-0.5\Delta^{1/2}$ were evaluated. These six approximating sequences and the exact solution $\tilde{x}$ are illustrated in Figure A in the (x,s) scale. Here $\tilde{x}_1 = \{\tilde{x}_n(0.01)\}$, $\tilde{x}_2 = \{\tilde{x}_n(0.0025)\}$, $\tilde{x}_3 = \{\tilde{x}_n(0.000625)\}$ and similarly $\tilde{x}^*_1 = \{\tilde{x}^*_n(0.01)\}$, $\tilde{x}^*_2 = \{\tilde{x}^*_n(0.0025)\}$, $\tilde{x}^*_3 = \{\tilde{x}^*_n(0.000625)\}$. This figure clearly illustrates that for this particular problem whereas the approximations provided by $\tilde{x}_1$, $\tilde{x}_2$ and $\tilde{x}_3$ are quite crude, the approximations provided by $\tilde{x}^*_1$, $\tilde{x}^*_2$ and $\tilde{x}^*_3$, particularly both $\tilde{x}^*_2$ and $\tilde{x}^*_3$, are exceptionally accurate, being virtually indistinguishable from each other and from the exact solution. The fact that $\tilde{x}^*_1$ is not too accurate reflects the fact that when using these approximations, one must begin with a reasonably small value of Δ . Indeed if we note that $\tilde{x}^*_1$ results from correcting the boundary $\tilde{x}_1$ which is obtained by approximating the Wiener process on the interval $[0,1]$ by a simple random walk involving only 100 time steps, it is somewhat amazing that $\tilde{x}_1^*$ achieves the accuracy that it does.

Figure A

The exceptional performance of the refined approximation (2.9) in Van Moerbeke's problem leads us to hope that the same type of behavior will occur in the one-armed bandit problem. In order to examine this possibility, the one-armed bandit problem was approximated by three different analogous discrete problems, those corresponding to $\Delta = 1.0$, 0.25 and 0.0625. For each fixed value of Δ the computations were carried out for the region $1 \leq s \leq 100$ and for all grids specified by values of the parameter c varying from 0 to $\Delta^{1/2}$ in steps of 0.01. Thus each individual member of each of the three sequences $\{\tilde{y}_n(\Delta)\}$ is located to within an error of 0.01. In addition the corrected sequences $\{\tilde{y}_n^*(\Delta)\}$ defined by $\tilde{y}_n^*(\Delta) = \tilde{y}_n(\Delta) - 0.5\Delta^{1/2}$ were evaluated. These six approximating sequences are illustrated in Figure B. Here $\tilde{y}_1 = \{\tilde{y}_n(1.0)\}$, $\tilde{y}_2 = \{\tilde{y}_n(0.25)\}$, $\tilde{y}_3 = \{\tilde{y}_n(0.0625)\}$ and similarly $\tilde{y}_1^* = \{\tilde{y}_n^*(1.0)\}$, $\tilde{y}_2^* = \{\tilde{y}_n^*(0.25)\}$, $\tilde{y}_3^* = \{\tilde{y}_n^*(0.0625)\}$. This figure clearly indicates that for the one-armed bandit problem the approximations provided by $\tilde{y}_1^*$, $\tilde{y}_2^*$ and y_3^* are exceptionally accurate, these curves being indistinguishable from one another.

Figure B

As pointed out in the introduction, for applications of the solution of the one-armed bandit problem, the (z,t) scale where $z = y/s$ and

$t = 1/s$ is more appropriate. In order to obtain an accurate approximation to the optimal stopping boundary in the (z,t) scale in as efficient a manner as possible, the computations were carried out as follows: Beginning with a very small value of Δ, the boundary was approximated in a small interval of s in the manner described above. Successively larger values of Δ were then employed to approximate the boundary in successively larger intervals of s. These approximations to the optimal boundary, determined in overlapping intervals of s, were then superimposed to obtain the final approximation to the optimal boundary. Since the values of Δ used were chosen in such a way as to yield the desired accuracy, only the value $c = 0$ was used in these computations. The computations were carried out using both the direct and the truncation method. The truncation method reduced the computation time required by a factor of approximately two. The resulting approximation to the optimal stopping boundary is illustrated in Figure C together with the asymptotic expansions of Chernoff and Ray (1965). Here $\tilde{z}_0$ and $\tilde{z}_1$ denote the boundaries obtained using the asymptotic expansions for t close to 0 (s large) and t close to 1 (s close to 1) respectively and z denotes the boundary obtained by means of the computations described above.

Figure C

---

4. DISCUSSION

As indicated in the introduction, the one-armed bandit problem has arisen in a number of statistical applications and consequently considerable effort has been devoted to obtaining approximations to the optimal stopping boundary. Chernoff and Ray (1965) first formulated the problem and were able to derive asymptotic expansions for this boundary. In a later paper, Chernoff (1967) presents a small table of this boundary. Although it is not indicated in the paper, this approximation was obtained by approximating the Wiener process by a sum of independent normal random variables and applying a backward induction to the approximating process (private communication from Herman Chernoff). In carrying out this backward induction the normal distribution was approximated by a discrete distribution thus allowing the integrations involved in each stage of the backward induction algorithm to be replaced by summations. Mallik (1971) presents a more detailed table of this boundary and indicates without clarification that a refined version of the technique used by Chernoff was used to obtain his table. Another detailed table appears in Chernoff (1971 and 1972) and it is in these references that it is first suggested that the $Y'(s')$ process of Section 2 be used to obtain an approximation to the optimal boundary for the one-armed bandit problem.

The purpose of the present paper was to describe explicitly how this approximation could be carried out and to demonstrate that by the use of the "correction for continuity" given in (2.9) this approximation could be made exceptionally accurate in an efficient manner. Obtaining the present approximation to the boundary of the one-armed bandit problem involved ten separate runs, the i-th run approximating the boundary in the region $s = 1$ to $s = 1 + 1550 \cdot \Delta_i$ using a grid determined by $c = 0$ and $\Delta_i = 10^{-4} \cdot 4^{i-2}$. The entire computation, approximating the boundary in the region $1 \leq s \leq 10{,}000$ required just 36 seconds of computation time on the IBM 370/168 at UBC. The objective in the present computation was to obtain an accurate approximation in the (z,t) scale. Detailed examination of the computer output leads to the empirical estimate that in the (z,t) scale, for each fixed value of t, the boundary has been located to within an error of 0.001.

For the one-armed bandit problem it is convenient to think of

$$\alpha = y/s^{1/2} = y^*/s^{*1/2} \tag{4.1}$$

as the number of standard deviations that the current estimate of μ is away from 0, and of

$$\beta = \Phi(\alpha) \tag{4.2}$$

(Φ denotes the standard normal cumulative) as the significance of the data for rejecting the hypothesis $\mu = 0$ in favor of the alternative that $\mu < 0$. Note that $t = 1/s$ is the fraction of the total available information already collected. Then the optimal procedure can be regarded as stopping as soon as the hypothesis $\mu = 0$ is rejected in favor of the alternative $\mu < 0$ at a <u>nominal significance level</u>.

$$\tilde{\beta}(t) = \Phi(\tilde{\alpha}(t)) = \Phi(\tilde{y}(s)/s^{1/2}) \tag{4.3}$$

which varies with t. In variations of the one-armed bandit problem in which the data is not normally distributed, it seems reasonable to use this nominal significance level as a stopping criterion. In order to facilitate future use of the results of this paper, the optimal boundaries $\tilde{z}(t)$, $\tilde{\alpha}(t)$ and $\tilde{\beta}(t)$ are presented in Table 1.

Table 1

As indicated in the introduction, these same methods could be applied with equal facility to any optimal stopping problem of the general form described there. In Petkau (1977) these methods have been employed to obtain the optimal continuation region for a stopping problem in which the optimal continuation region can be described as the set $\{(y,s): \tilde{y}_1(s) < y < \tilde{y}_2(s),\ s > 1\}$ where $\tilde{y}_1(s) \neq \tilde{y}_2(s)$.

The connection between such optimal stopping problems and free boundary problems involving the heat equation of the form (1.4) makes it clear that these same methods could be used to determine numerical solutions of such free boundary problems. The problem of obtaining numerical solutions to free boundary problems has received considerable attention in the literature (see, for example, Sackett (1971) and Meyer (1977)). Whether the methods proposed here provide a reasonable alternative to these more general methods is a question that remains to be answered.

Acknowledgment

The author would like to acknowledge several helpful conversations with Herman Chernoff.

Table 1. One-Armed Bandit Boundary

t	$\tilde{z}(t)$	$\tilde{\alpha}(t)$	$\tilde{\beta}(t)$
0.0001	-0.0353	-3.5318	0.00021
0.0002	-0.0473	-3.3431	0.00041
0.0003	-0.0559	-3.2275	0.00062
0.0004	-0.0630	-3.1487	0.00082
0.0005	-0.0689	-3.0830	0.00102
0.0006	-0.0742	-3.0285	0.00123
0.0007	-0.0789	-2.9821	0.00143
0.0008	-0.0832	-2.9416	0.00163
0.0009	-0.0872	-2.9059	0.00183
0.0010	-0.0909	-2.8731	0.00203
0.0020	-0.1187	-2.6547	0.00397
0.0030	-0.1381	-2.5207	0.00586
0.0040	-0.1534	-2.4262	0.00763
0.0050	-0.1661	-2.3486	0.00942
0.0060	-0.1771	-2.2859	0.01113
0.0070	-0.1869	-2.2338	0.01275
0.0080	-0.1955	-2.1862	0.01440
0.0090	-0.2035	-2.1446	0.01599
0.0100	-0.2107	-2.1073	0.01754
0.0200	-0.2624	-1.8555	0.03176
0.0300	-0.2955	-1.7062	0.04398
0.0400	-0.3196	-1.5980	0.05502
0.0500	-0.3385	-1.5138	0.06504
0.0600	-0.3537	-1.4441	0.07436
0.0700	-0.3665	-1.3854	0.08297
0.0800	-0.3771	-1.3333	0.09121
0.0900	-0.3866	-1.2887	0.09875
0.1000	-0.3944	-1.2472	0.10616
0.1200	-0.4074	-1.1762	0.11976
0.1400	-0.4174	-1.1155	0.13232
0.1600	-0.4252	-1.0630	0.14390
0.1800	-0.4310	-1.0159	0.15484
0.2000	-0.4354	-0.9735	0.16516
0.2500	-0.4413	-0.8827	0.18871
0.3000	-0.4415	-0.8061	0.21008
0.3500	-0.4378	-0.7400	0.22964
0.4000	-0.4308	-0.6812	0.24788
0.4500	-0.4207	-0.6272	0.26527
0.5000	-0.4083	-0.5774	0.28184
0.5500	-0.3938	-0.5310	0.29771
0.6000	-0.3764	-0.4860	0.31350
0.6500	-0.3560	-0.4416	0.32939
0.7000	-0.3333	-0.3983	0.34520
0.7500	-0.3074	-0.3549	0.36133
0.8000	-0.2774	-0.3102	0.37821
0.8500	-0.2423	-0.2628	0.39634
0.8700	-0.2263	-0.2426	0.40417
0.8900	-0.2088	-0.2213	0.41243
0.9100	-0.1894	-0.1986	0.42130
0.9300	-0.1676	-0.1738	0.43102
0.9500	-0.1421	-0.1457	0.44206
0.9600	-0.1272	-0.1298	0.44835
0.9700	-0.1103	-0.1120	0.45542
0.9800	-0.0902	-0.0911	0.46369
0.9900	-0.0639	-0.0642	0.47440
0.9950	-0.0452	-0.0453	0.48194
1.0000	0.0	0.0	0.50000

BEST AVAILABLE COPY

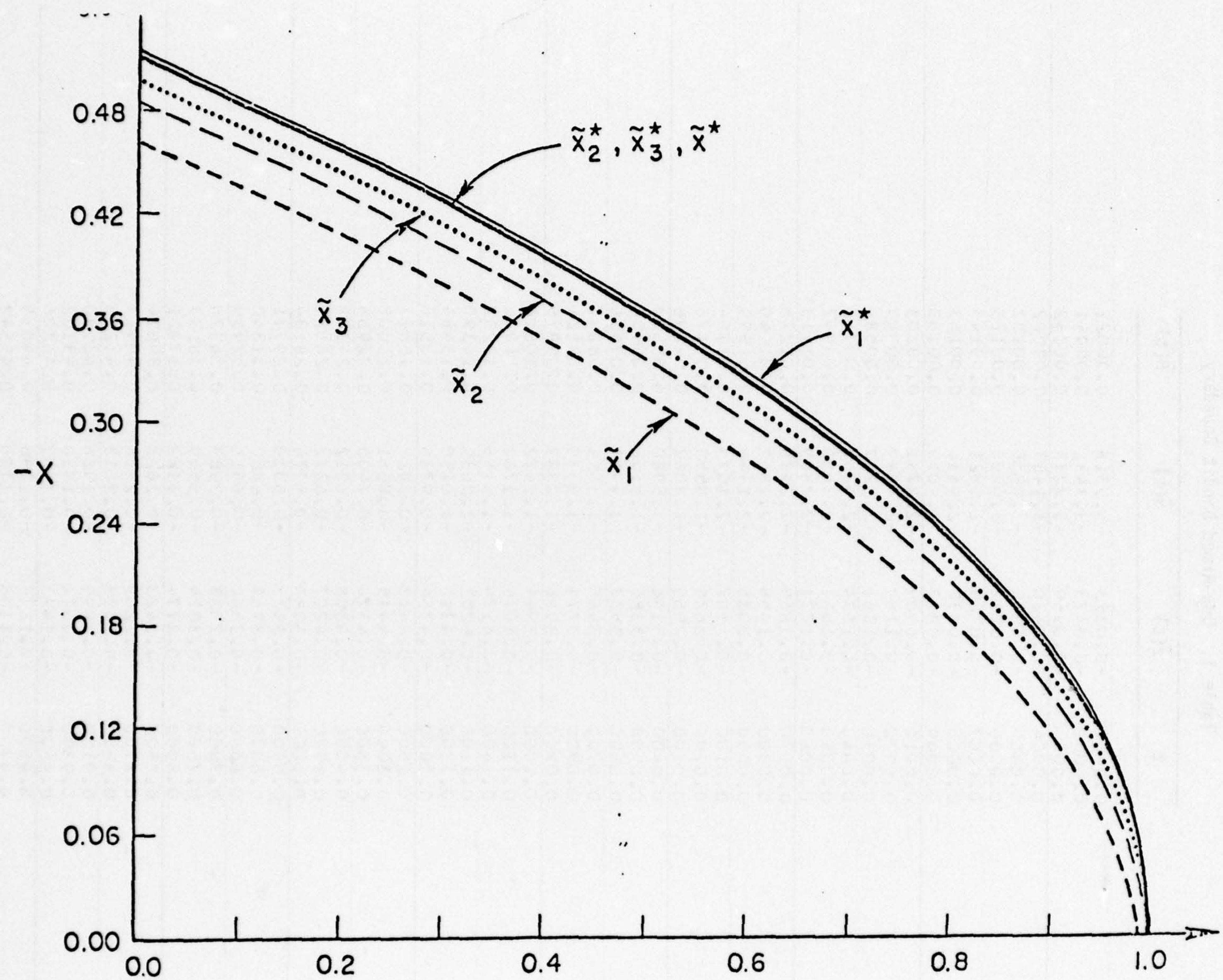


Figure A Van Moerbeke's Problem

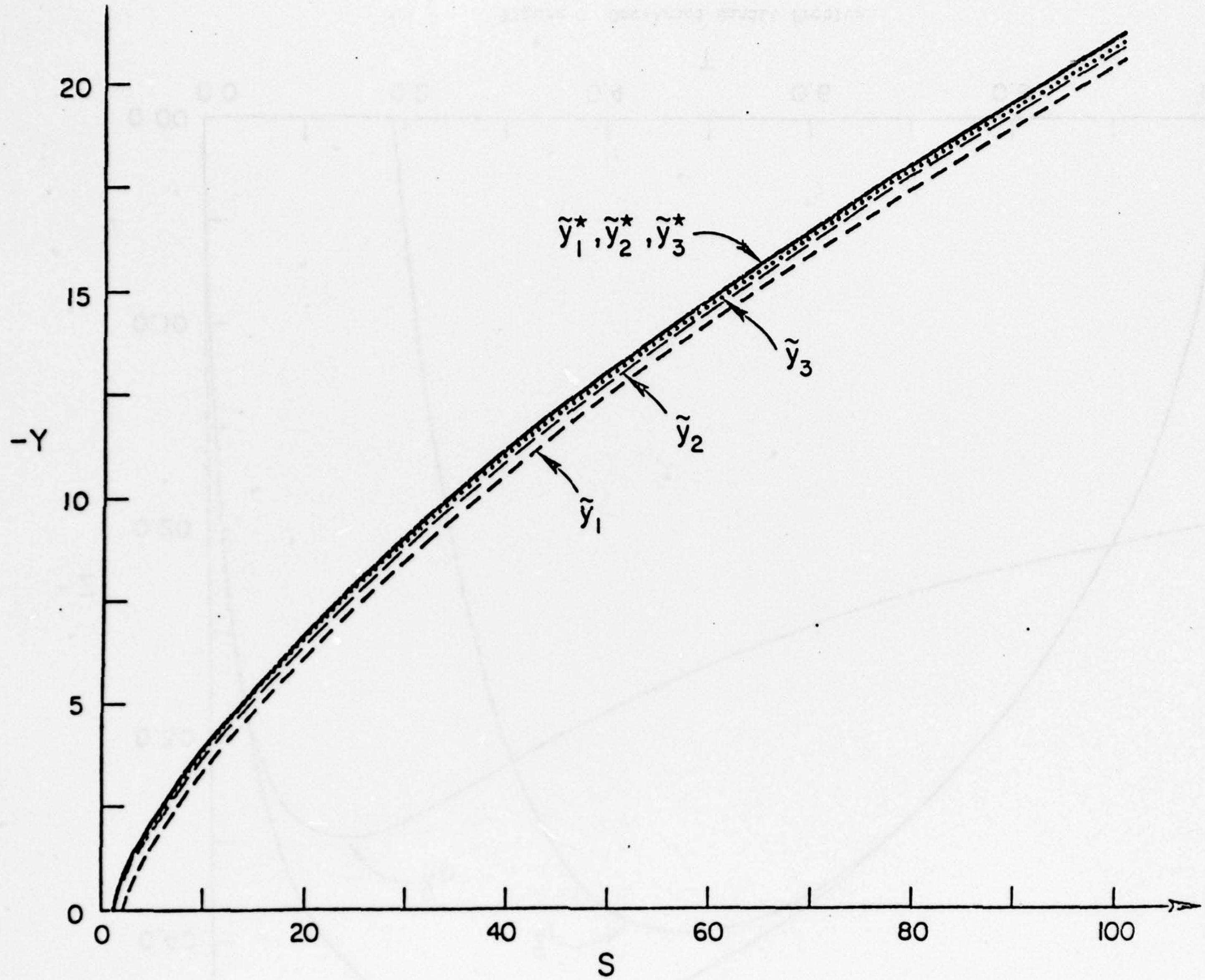


Figure B One-Armed Bandit Problem

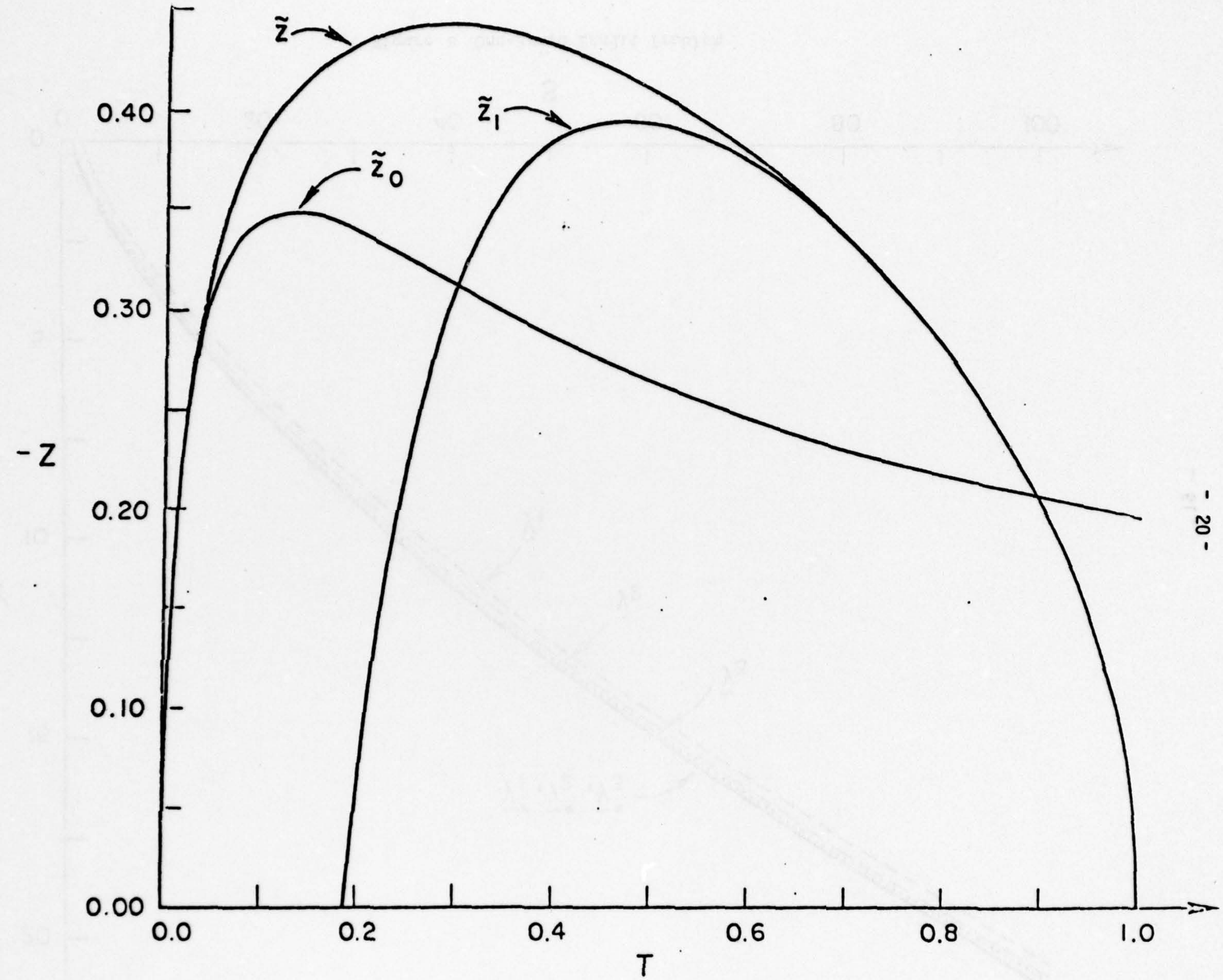


Figure C One-Armed Bandit Problem

REFERENCES

Chernoff, H. (1967), "Sequential models for clinical trials", Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, IV, 805-812.

_____ (1968), "Optimal stochastic control", Sankhyā, Series A, 30, 221-252.

_____ (1971), The efficient estimation of a parameter measurable by two instruments of unknown precisions", Optimizing Methods in Statistics, J.S. Rustagi, ed. Academic Press, New York.

_____ (1972), Sequential Analysis and Optimal Design, SIAM monograph, Philadelphia.

_____ and Petkau, A.J. (1976), "An optimal stopping problem for sums of dichotomous random variables", Annals of Probability, 4, 875-889.

_____ and Ray, S.N. (1965), "A Bayes sequential sampling inspection plan", Annals of Mathematical Statistics, 36, 1387-1407.

Feller, W. (1968), An Introduction to Probability Theory and its Applications, Vol. 1, 3rd edition, John Wiley and Sons, Inc., New York.

Mallik, A.K. (1971), "Sequential estimation of the common mean of two normal populations", Technical Report No. 42, Department of Statistics, Stanford University.

Meyer, G.H. (1977), "One-dimensional parabolic free boundary problems", SIAM Review, 19, 17-34.

Petkau, A.J. (1977), "Sequential medical trials for comparing an experimental with a standard treatment", Journal of the American Statistical Association (to appear).

Sackett, G.G. (1971), "Numerical solution of a parabolic free boundary problem arising in statistical decision theory", Mathematics of Computation, 25, 425-434.

Van Moerbeke, P.L.J. (1974), "Optimal stopping and free boundary problems," Rocky Mountain Journal of Mathematics, 4, 539-578.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 8	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Approximate Solutions for Certain Optimal Stopping Problems		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) A. John Petkau		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0555
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mathematics University of British Columbia Vancouver, CANADA		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-331)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE December 21, 1977
		13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

optimal stopping, Wiener process, one-armed bandit, backward induction

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(see reverse side)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

This paper presents simple approximate methods which lead to arbitrarily accurate numerical approximations to the optimal continuation regions of optimal stopping problems involving a zero drift standard Wiener process. The methods involve approximating the Wiener process by a simple random walk, solving the analogous problem for the random walk and subsequently applying a "correction for continuity" due to Chernoff and Petkau to the solution of the discrete problem. The methods are illustrated in a problem due to Van Moerbeke for which the exact solution is known and in the one-armed bandit problem which has arisen in several different contexts in the statistical literature.